

# Utilisation de réseaux neuromimétiques pour la localisation automatique d'événements microsismiques

Use of neuromimetic networks for the automatic location of microseismic events  
Die Verwendung von auf Informationssystemen basierenden Methoden zur Überwachung seismischer Aktivitäten, die durch Abbau verursacht wurden

P. BIGARRÉ, S. LABLÉE, A. DODO AMADOU, J. P. PIGUET & J. P. JOSIEN, INERIS, Ecole des Mines, Nancy, France

**RESUME :** La télésurveillance microsismique en temps réel permet de caractériser la réponse du massif rocheux, par l'analyse appropriée des sismogrammes enregistrés. Parmi les paramètres estimés, la localisation (X,Y,Z,T) des ruptures détectées peut être considérée comme l'information la plus fondamentale. La précision de l'estimation dépend étroitement de la justesse avec laquelle les pointés d'arrivée d'onde aux différentes stations sont effectués. Dans cet article, nous présentons le développement d'un réseau de neurones permettant le pointé automatique de signaux sismiques, couplé à un algorithme «classique» du type STA/LTA. Les résultats de cette méthode couplée sont présentés, ainsi que l'intégration de l'outil sur le site minier de Gardanne. Les perspectives d'applications de réseaux de neurones sont discutées.

**ABSTRACT :** Real-time microseismic monitoring applied to mine-induced seismicity permits to characterize rock mass behaviour, considering appropriate, fast, automatic analysis of recorded waveforms. Among usual calculated parameters for each tremor, hypocenter location (X,Y,Z,T) is of the most importance, whereas its accuracy depends straightly from quality of first wave arrival-time pick-up. In this article, we present integration of neural nets technique to automatic first wave time estimation, parallelly processed with conventional STA/LTA. Results of this coupled method, site integration and use, as well as future prospects are discussed.

**ZUSAMMENFASSUNG :** Die microseismische Echtzeitüberwachung ermöglicht durch eine sachgemäße Untersuchung der registrierten Seismogramme eine genaue Bestimmung des Verhaltens des Gebirges. Die Lokalisierung (X, Y, Z, T) der wahrgenommenen Brüche ist der bedeutenste der berechneten Parameter. Die Genauigkeit der Berechnung dieser Lokalisierung steht in sehr engem Zusammenhang mit der Präzision mit der die Erdbebenwellen von den verschiedenen Aufnahmestationen erfasst werden. In diesem Aufsatz beschreiben wir die Entwicklung eines neuronalen Netzes zur Erfassung dieser Wellen, welches mit einem "klassischen" Algorithmus der Art STA/LTA gekoppelt ist. Weiter werden die Resultate dieser Anwendung, sowie die Integration des Systems auf dem Bergwerk in Gardanne erörtert.

## Introduction

En 1992, l'Unité d'Exploitation de Provence des HBCM, en collaboration avec l'INERIS et le LDG<sup>(\*)</sup>, a mis en oeuvre un réseau de télésurveillance microsismique composé de 9 stations de type géophones. Ce réseau permet d'enregistrer et d'analyser l'activité microsismique des chantiers situés à 1100 mètres de profondeur environ, et exploités par longue taille foudroyée. La station d'acquisition détecte, par scrutation et dépassement de seuil sur au moins 3 voies, les ruptures de terrains et enregistre les signaux pendant 30 secondes à une fréquence d'échantillonnage de 155 Hz. En moyenne, ce sont 50 à 100 événements quotidiens de magnitude 1.5 à 3 sur l'échelle de magnitude locale qui sont détectés, traités et stockés (voir figure 1).

L'abondance d'événements et donc de l'information à collecter à partir des sismogrammes pour l'interprétation du comportement du massif rocheux a nécessité le développement d'outils de traitement automatique. En particulier, l'ensemble des signaux est traité par une procédure de détection d'événement permettant le pointage précis, à la valeur du pas d'échantillonnage près, de chacun des 9 temps d'arrivée  $t_i$  des ondes (méthode basée sur un algorithme de type STA/LTA). L'application d'une méthode numérique inverse permet alors d'estimer la localisation de l'événement dans l'espace.

Cette procédure classique de détermination du vecteur  $(t_i)_{i=1,9}$  reste

néanmoins peu robuste pour les signaux présentant certaines singularités comme :

- un rapport signal/bruit faible,
- plusieurs signaux distincts sur le même enregistrement,
- des bruits parasites divers (bips, créneaux) (voir figure 2).

L'existence d'un pourcentage significatif d'enregistrements mal localisés par procédure automatique se double d'un autre problème : celui de leur identification rapide par une voie autre que celle du traitement dit interactif ou manuel, que le nombre d'événements à vérifier rend parfois impossible.

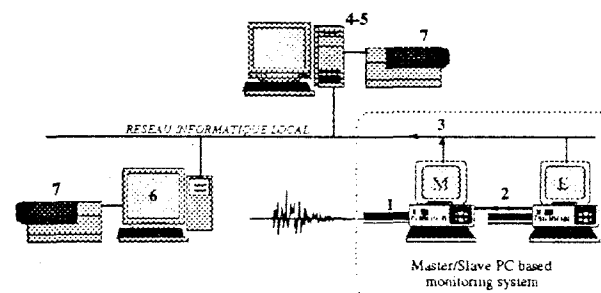


Figure 1 : Système automatique d'acquisition et de traitement des données opérationnel à l'U.E. Provence: 1) événement détecté 2) acquisition, passage en mode esclave, prétraitement des données 3) transmission sur machine multitâche 4) pointage d'onde, localisation et traitement complet 5) archivage automatique 6) traitement interactif / validation des données 7) édition de rapports

(\*) HBCM : Houillères du Bassin Centre Midi (Charbonnage de France)  
INERIS : Institut National de l'Environnement Industriel et des Risques  
LDG : Laboratoire de Détection et de Géophysique

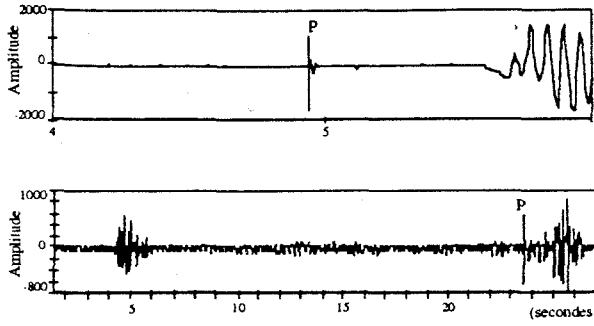


Figure 2 : deux exemples classiques de pointés défectueux : 1) pointage d'un bip parasite non filtré - 2) pointage d'un second événement sur le même signal. Dans les deux cas, le résultat de localisation (X,Y,Z,T) peut s'en retrouver très perturbé, à l'échelle du réseau.

Dans le cadre de ses études conduites sur la sismicité induite par exploitation minière, l'INERIS a développé et testé un logiciel de pointage automatique de temps d'arrivée d'ondes basé sur la technique des réseaux neuromimétiques (SCOP : Système Connexionniste de Pointage automatique de début d'onde). Ce logiciel a pour objectif de répondre aux deux exigences précitées, à savoir :

- 1- fiabiliser le traitement de localisation en diminuant le nombre d'événements pointés de manière trop peu satisfaisante,
- 2- mettre en oeuvre un calcul de «coefficient de confiance» permettant d'identifier automatiquement les dits événements, afin de minimiser les tâches de vérification et de traitement manuels.

Les principaux résultats sont présentés ci-après, ainsi que les perspectives d'applications et d'intégration dans le dépouillement automatique.

### Le paradigme connexionniste

Le modèle formel du neurone décrit l'entité neurale comme un automate recevant des valeurs en entrée et y appliquant une certaine fonction de transfert pour calculer une valeur de sortie [Mc Culloch, Pitts, 1943] (voir figure 3).

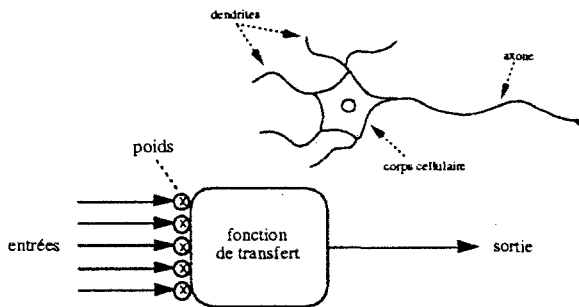


Figure 3 : schéma du modèle formel de neurone: les entrées modélisent les dendrites, la sortie l'axone.

Les neurones formels sont interconnectés de manière similaire aux neurones biologiques, les connexions étant alors pondérées. La valeur de sortie d'un neurone est calculée par application d'une fonction de transfert à la somme pondérée de ses entrées.

Ainsi, si  $\varphi$  est la fonction de transfert d'un neurone,  $(e_i)_{i=1,n}$  les valeurs de ses entrées, et  $(\mu_i)_{i=1,n}$  les poids synaptiques de ses entrées, la valeur de sortie  $s$  de ce neurone sera définie par :

$$s = \varphi\left(\sum_{i=1}^n \mu_i e_i\right).$$

Les sorties du réseau neuromimétique, correspondant aux valeurs

de sortie des neurones dont l'axone est libre, sont calculées par application successive de cette règle à chaque neurone. L'ordre d'application diffère suivant les types de réseaux. Dans le cas du classique réseau «perceptron multicouche» [Rosenblatt 1958], où les neurones sont groupés en couches successives, le calcul s'effectue de couche en couche, de l'entrée vers la sortie. Quant aux entrées du réseau, elles correspondent aux neurones dont les dendrites sont libres.

En 1949, Hebb émet l'hypothèse que le cerveau s'adapte à son environnement par modification des connexions synaptiques [Hebb 1949]. Par analogie, le réseau neuromimétique réalise son apprentissage par modification des poids synaptiques de chaque neurone. Deux types d'apprentissage sont distingués : l'apprentissage *non supervisé*, dans lequel le réseau voit itérativement ses poids modifiés pour converger de lui-même vers un certain état d'équilibre, et l'apprentissage *supervisé*, dans lequel l'objectif est d'ajuster les poids pour minimiser la différence entre les sorties effectives du réseau et des sorties pré-définies.

Le système SCOP utilise ce second type d'apprentissage, en modifiant itérativement les poids synaptiques après présentation d'exemples (couples entrées-sorties désirées), pour minimiser l'écart. Après apprentissage, le réseau peut délivrer des valeurs de sortie pour des entrées non apprises. Pratiquement, le principe de base de l'algorithme d'apprentissage consiste à minimiser une fonction caractérisant l'erreur en sortie par la méthode des moindres carrés (D.Parker 1985, Y.LeCun 1986, D.E.Rumelhart et al 1986). Chaque connexion est donc modifiée en fonction de sa participation à l'erreur totale, donc en fonction de la dérivée de la fonction d'erreur  $E$  (fonction des sorties et des valeurs cibles, posée en général comme étant la somme sur les neurones de sortie des différences entre sortie effective et sortie désirée) par rapport à cette connexion :

$$\Delta\mu_{i,j} = -\varepsilon \frac{\partial E}{\partial \mu_{i,j}}$$

(où  $\varepsilon$  est un coefficient arbitraire et  $\mu_{i,j}$  la valeur de la connexion du neurone  $j$  de la couche  $i$  à son successeur). Si l'on pose  $H_i = \sum_j \mu_{i,j} x_j$ , où  $x_j$  est la valeur du neurone  $j$  de la couche  $i-1$ , si  $y_i$  est la sortie du neurone  $i$ , et si  $\varphi$  est la fonction d'activation des neurones ( $y_i = \varphi(H_i)$ ) la dérivée partielle se décompose alors pour la dernière couche en termes directement mesurables :

$$\frac{\partial E}{\partial \mu_{i,j}} = \frac{\partial E}{\partial y_i} \frac{\partial y_i}{\partial H_i} \frac{\partial H_i}{\partial \mu_{i,j}} = (y_i - y_i^{\text{désiré}}) \cdot \varphi'(H_i) \cdot x_j$$

La rétro-propagation consiste alors à propager l'erreur de la dernière couche (calculée ci-dessus) vers les couches précédentes :

$$\frac{\partial E}{\partial \mu_{i,j}} = \sum_k \left[ \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial H_k} \frac{\partial H_k}{\partial \mu_{i,j}} \right] = \sum_k \left[ \frac{\partial E}{\partial y_k} \cdot \varphi'(H_k) \cdot y_i \right]$$

Les  $\frac{\partial E}{\partial y_i}$  pour les neurones d'une couche se calculent à partir des

neurones  $\frac{\partial E}{\partial y_k}$  calculés pour les neurones des couches précédentes.

En pratique, l'apprentissage par rétro-propagation se fait en deux phases. Durant la première phase, un exemple est présenté en entrée et propagé à travers le réseau de couche en couche. Une valeur de sortie est obtenue pour chaque neurone de la dernière couche. Cette valeur est ensuite comparée à la valeur attendue, ce qui fournit un vecteur d'erreur :

$$\begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix} = \begin{pmatrix} \text{sortie neurone 1} - \text{sortie attendue neurone 1} \\ \vdots \\ \text{sortie neurone n} - \text{sortie attendue neurone n} \end{pmatrix}$$

La seconde phase consiste alors à rétro-propager ce vecteur dans le réseau (grâce aux formules décrites ci-dessus) afin de modifier successivement les poids des connexions des neurones de chaque couche.

### Le connexionnisme dans SCOP

SCOP utilise un réseau neuromimétique afin de calculer le temps d'arrivée des ondes P au sein du signal sismique enregistré. Une représentation de ce signal est donc constituée par les valeurs d'entrée du réseau, la valeur de sortie étant le temps séparant le début du signal du début des ondes P dans ce signal. SCOP est organisé autour d'un modèle du type : Cascade-Corrélation, décrit en 1991 par [Fahlman 1991]. Il s'agit d'un modèle dit «ontogénétique» (\*), créant les neurones dont il a besoin au fur et à mesure de son apprentissage. Les avantages principaux de ce type de modèle sont la mise en place d'une morphologie de structure adaptée au problème à résoudre, et une rapidité d'apprentissage accrue. L'algorithme de base utilisé pour l'apprentissage est une variante de la rétro-propagation, l'algorithme «Quickprop» utilisant une méthode du second ordre (se rapportant à la méthode de Newton) au lieu de la simple descente de gradient explicitée plus haut.

La méthode consiste ensuite, partant d'un nombre minimal de neurones, à entraîner le réseau jusqu'à ce que l'erreur commise soit inférieure à un seuil prédéfini. Si, au cours de l'apprentissage, l'erreur ne décroît plus (stagnation), un nouveau neurone est ajouté au noeud du réseau où sa présence a une influence maximale sur le résultat (corrélation maximale entre l'erreur due à ce neurone et l'erreur totale du réseau). Un nombre maximal de neurones ajoutés est fixé au préalable afin d'éviter un apprentissage infini sans succès.

### Implantation informatique

SCOP a été développé en langage C (norme ANSI) sous système UNIX. L'étude des performances de SCOP s'est appuyée à la fois sur des modules de tests écrits en C et sur le logiciel de calcul formel Mathematica® (Wolfram Research, Inc).

### Apprentissage

Afin de créer une base d'exemples de signaux pour l'apprentissage et le test de SCOP, 5000 signaux enregistrés sur le site de Gardanne par différents capteurs sur une période d'un an ont été soumis à un pointage par opérateur des temps d'arrivée d'ondes P, mettant à disposition du système une banque de signaux pointés de la manière la plus précise possible. 3000 signaux ont été utilisés pour l'apprentissage, les 500 signaux restant étant destinés à tester le comportement du réseau vis-à-vis de signaux non appris.

### Prétraitement-Entrées

La complexité des formes des signaux sismiques impose un prétraitement précédant leur présentation en entrée du réseau neuromimétique. Ce prétraitement a pour but de générer une représentation «parlante» des données, de laquelle le réseau puisse aisément retirer des informations. La méthode retenue est celle dite de Gabor, ou transformée de Fourier mobile. Elle consiste à effectuer la transformée de Fourier de parties juxtaposées du signal. Un signal décomposé en N fenêtres donne alors lieu à N transformées, chacune exprimant le spectre de la fenêtre correspondante, et par extension l'image spectrale du signal à un certain instant. On obtient donc par cette série de spectres un diagramme fréquence-temps du signal sismique (Fig.4).

(\*) ontogénétique (ou ontogénique) : du terme ontogénèse, caractérisant la capacité d'un système complexe, ou organisme, à s'organiser et croître de lui-même.

C'est ce diagramme fréquence-temps qui est livré en entrée au réseau neuromimétique, sous la forme d'une matrice NxM, où M est la taille en points de la transformée.

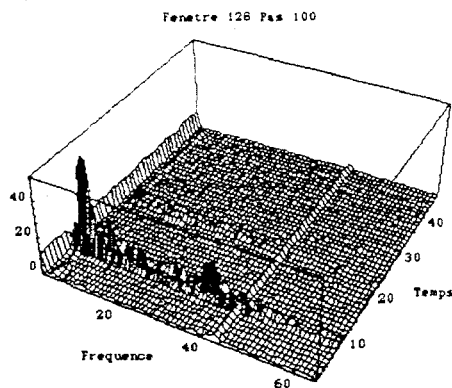


Figure 4 : Diagramme fréquence temps obtenu par FFT mobile (méthode de Gabor) à partir d'un signal sismique.

### Sorties

Nous avons choisi de rechercher une «plage de pointage» (i.e. un intervalle  $[t_1, t_2]$  contenant le début des ondes P) plutôt qu'un pointage exact  $t$  : le signal est découpé en une série de plages temporelles de dimension fixée, chacune étant mise en correspondance avec un neurone de sortie du réseau. Le neurone de sortie ayant la plus forte valeur indique la plage devant contenir le début des ondes P.

Ainsi, pour un pointage dans l'intervalle total [1 s., 31 s.] découpé en 60 plages de 0,5 seconde, la plage [ 2,5 s., 3 s.] est représentée en sortie par le vecteur (0 0 0 1 0 0 ... 0 0 0) (tous les neurones de sortie à 0 excepté le 4ème). Mais puisqu'en pratique les sorties ne sont pas binaires mais réelles dans  $[0, 1]$ , la plage en question pourra être représentée par n'importe quel vecteur de sortie dont la composante maximale est la 4ème.

Cette méthode a l'avantage de discrétiser les sorties théoriques du réseau, ce qui a pour effet d'augmenter ses performances. En effet, les réseaux de neurones formels s'utilisent de manière naturelle en classificateurs, et les résultats sont généralement satisfaisants. De plus, le réseau est dans ce cas plus robuste : la fenêtre de pointage retenue étant celle qui correspond au neurone de sortie maximale, une légère erreur sur chaque neurone de sortie a peu d'influence sur le résultat final.

### Résultats de SCOP utilisé seul

Le réseau apprend en moins d'une centaine d'itérations, c'est-à-dire que l'erreur commise par le réseau sur le pointage des signaux réservés au test (signaux non appris) est minimale après avoir présenté moins de 100 fois l'ensemble des signaux d'apprentissage prétraités au réseau et modifié les connexions en fonction de l'erreur. Le nombre de neurones est alors de 410 environ (environ 400 neurones d'entrée, ce chiffre dépendant des paramètres du prétraitement, 4 neurones de sortie, et en moyenne 5 neurones cachés créés par l'algorithme au fur et à mesure des besoins de l'apprentissage).

Après cette phase d'apprentissage, la réponse de SCOP a été testée sur 500 signaux non appris. Le tableau suivant résume les pourcentages de réponses correctes (i.e. pour lesquelles le début des ondes P était bien à l'intérieur de la fenêtre délivrée par SCOP). Ces résultats sont fonction de la taille de la fenêtre de pointage ( $t_2 - t_1$  pour une fenêtre  $[t_1, t_2]$ ).

Test sur les signaux appris (3000 signaux) :

Taille de la fenêtre de pointage	Résultats avec l'algorithme classique	Résultats avec SCOP
0,8 seconde	80,0 %	100 %
1,3 seconde	83,3 %	100 %

Test sur les signaux non appris (500 signaux) :

Taille de la fenêtre de pointage	Résultats avec l'algorithme classique	Résultats avec SCOP
0,8 seconde	80,0 %	82,5 %
1,3 seconde	83,3 %	89,8 %

Ces résultats appellent les commentaires suivants :

- l'égalité des résultats de l'algorithme classique dans les deux cas (signaux appris, signaux non appris) provient de la sélection aléatoire des signaux de test au sein de l'ensemble des signaux, l'erreur étant alors statistiquement répartie de manière homogène sur les deux classes (appris, non appris),
- l'apprentissage est réussi, puisque SCOP ne fait aucune erreur de pointage sur les signaux appris,
- les performances augmentent avec la taille de la fenêtre de pointage, ce qui est normal puisque l'on demande alors une précision moindre,
- les performances de SCOP sont en général supérieures aux performances de l'algorithme dit «classique» utilisé jusqu'à présent (basé sur STA/LTA), mais l'écart entre les deux méthodes est relativement faible pour des fenêtres de pointage suffisamment précises (0,8 seconde, voire 0,65 seconde où les tests montrent même une légère supériorité de l'algorithme classique).

#### Un système de pointage automatique mixte

L'étude de la répartition des erreurs de SCOP et de l'algorithme de STA/LTA montre que le taux de recouvrement des erreurs des deux systèmes ne dépasse pas 3 à 4 % (toujours pour un pointage à 0,65 seconde près). Cela signifie que les deux systèmes sont simultanément dans l'erreur dans moins de 4 % des cas. Par conséquent, l'un au moins des deux systèmes pointe correctement dans plus de 96 % des cas. Nous avons donc recherché un mode de coopération des deux systèmes permettant de mettre en évidence la méthode offrant le pointage le plus pertinent pour chaque signal. Le principe en est le suivant : un pointage simultané par SCOP et STA/LTA du signal à traiter est effectué, le premier délivrant une plage temporelle  $[t_1, t_2]$  et le second une valeur temporelle  $t$ . Il s'agit alors de déterminer la manière dont SCOP corrobore les résultats de STA/LTA en examinant la position de  $t$  par rapport à l'intervalle  $[t_1, t_2]$  :

- si  $t$  est dans  $[t_1, t_2]$ , le pointage considéré comme exact est  $t$ . On affecte à ce résultat un coefficient de confiance compris entre 0 et 1. Dans le cas présent, ce coefficient (nommons-le  $c$ ) vaut  $1-\varepsilon$  ( $\varepsilon$  représentant le doute dû aux 3 à 4 % de signaux mal pointés par les deux méthodes).
- si  $t$  n'est pas dans  $[t_1, t_2]$ , on cherche alors à pointer par STA/LTA un temps  $t'$  à l'intérieur de  $[t_1, t_2]$  (c'est l'opération dite de «repointage»). Si  $t'$  est trouvé, le pointage considéré comme exact est  $t'$ , avec un coefficient de confiance  $c_1$  inférieur à  $1-\varepsilon$  (car un doute subsiste sur le pointage). Dans le cas contraire ( $t'$  non trouvé), le pointage considéré comme exact est le pointage initial, avec un coefficient de confiance  $c_2$  inférieur à  $c_1$ , donc à  $1-\varepsilon$ .  $c_1$  et  $c_2$  peuvent être définis par une fonction de la distance de  $t$  (resp.  $t'$ ) au milieu de  $[t_1, t_2]$  : plus  $t$  (resp.  $t'$ ) est éloigné du résultat de SCOP, plus le doute est grand.

Résultats comparés sur les 500 signaux de test (non appris) :

Taille de la fenêtre de pointage	Résultats avec l'algorithme classique	Résultats avec SCOP seul	Résultats avec le système mixte (*)
0,8 seconde	80,0 %	82,5 %	94,5 %
1,3 seconde	83,3 %	89,8 %	96,5 %

#### Conclusion et perspectives

L'utilisation de techniques connexionnistes pour l'aide à l'analyse de signaux sismiques s'est montrée possible et prometteuse. Mais, plus encore, l'utilisation simultanée de méthodes connexionnistes et de méthodes algorithmiques classiques au sein d'un même système mixte s'est révélée très performante, les résultats obtenus dépassant en qualité ceux de l'une ou l'autre des méthodes prises séparément. En outre, le système actuel délivre non seulement une valeur de pointage, mais aussi un indice d'appréciation de la précision de ce pointage : le «coefficient de confiance». Ainsi, il est désormais possible de déterminer parmi les événements traités lesquels doivent être repointés prioritairement par un opérateur. La sismicité induite par les chantiers de la mine repose essentiellement sur la qualité de ce résultat. Concernant les perspectives de développement, SCOP devrait évoluer selon deux voies : 1 - extension des modules pour transformer SCOP en générateur de réseaux neurones, d'une grande simplicité de mise en oeuvre ; 2 - recherche d'utilisation de SCOP sur la classification automatique des signaux, et notamment la détection des enregistrements comprenant deux événements sismiques (voir fig. 1)

#### REFERENCES :

- [Changeux 1983] Changeux, J.P. (1983) «L'homme neuronal», Fayard.
- [Dodo Amadou 1993] Dodo Amadou A., «Sismologie appliquée à l'étude du comportement de la roche. Mise en oeuvre de la méthode des doublets.» DEA de Génie Civil et Minier de l'Institut National Polytechnique de Lorraine, 1993.
- [Fahlman 1988] Fahlman, S.E. (1988) «Faster-Learning Variations on Back-Propagation : An Empirical Study» in *Proceedings of the 1988 Connectionist Models Summer School*, Morgan Kaufman.
- [Fahlman 1991] Fahlman, S.E. & Lebiere, C. (1991) «The Cascade Correlation Learning Architecture», rapport interne School of Computer Science - Carnegie Mellon University, Pittsburgh PA. USA
- [Grossman & Morlet 1984] Grossman, A. & Morlet, J. (1984) «Decomposition of Hardy Functions Into Square Integrable Wavelets of Constant Shape», in *SIAM J.Math. vol.15 pp.723-736*
- [Hebb 1949] Hebb, D. (1949) «The organisation of Behavior», Wiley
- [Mc Culloch, Pitts, 1943] Mc Culloch, W.S. & Pitts, W. (1943) «A logical calculus of the ideas immanent in nervous activity», in *Bulletin of Mathematical Biophysics n°9 pp.127-147*
- [Meyer 1992] Meyer, Y. (1992) «Les ondelettes. Algorithmes et applications», Armand Collin
- [PDP 1986] Mc Clelland, J.L. & Rumelhart D.E. & the PDP Research Group (1986) «Parallel Distributed Processing. Explorations in the Microstructure of Cognition», MIT Press/Bradford
- [Rosenblatt 1958] Rosenblatt, F. (1958) «The perceptron : a probabilistic model for information storage and organization in the brain» in *Psychological Review n°65 pp.386-408*
- [V. Allen 1978] V. Allen, R. (1978) «Automatic earthquake recognition and timing from single traces» in *Bulletin of Seismological Society of America. vol. 68. n° 5, pp 1521-1532.*